

# Repaso de Algebra Matricial

*Francisco A. Ramírez*

*Semestre 2019S2*

## Repaso de Algebra Matricial

### Vectores y Matrices

Para fines de facilitar y hacer más transparente a notación matemática del texto se hace uso del algebra lineal.

### Terminología

En este libro un **vector** es una columna de números, denotado por:

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

La **transpuesta** de un vector, denotado por  $a' = (a_1, a_2, \dots, a_n)$ , es una fila de números, algunas veces llamado vector fila.

Una **matriz** es un arreglo rectangular de números de dimensión  $n \times k$  que puede ser escrita como:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{pmatrix}$$

El primer índice del elemento  $a_{ij}$  se refiere a la fila  $i$ , y el segundo índice se refiere a la columna  $j$ . Denotando el vector de la columna  $j$  por  $a_j$ , se ve que  $A$  consiste de  $k$  vectores  $a_1$  a  $a_k$ , y puede ser denotada como:

$$A = (a_1 \quad a_2 \quad \cdots \quad a_k)$$

El símbolo  $'$  denota la transpuesta de una matriz o vector,

$$A' = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & & & \\ a_{1k} & a_{2k} & \cdots & a_{nk} \end{pmatrix}$$

Las columnas de  $A$  son las filas de  $A'$ , y viceversa. Una matriz es **cuadrada** si  $n = k$ . Una matriz  $A$  es **simétrica** si  $A = A'$ . Una matriz cuadrada es llamada matriz **diagonal** si  $a_{ij} = 0$  para todo  $i \neq j$ . Note que una matriz diagonal es simétrica por construcción. La **matriz identidad**  $I$  es una matriz diagonal con todos los elementos de la diagonal igual a uno.

## Manipulación de matrices

Si dos matrices o vectores tienen una misma dimensión, estos pueden ser *sumados* o *sustraidos*. Sean  $A$  y  $B$  dos matrices de dimensiones  $n \times k$  con elementos  $a_{ij}$  y  $b_{ij}$ , respectivamente. Entonces  $A + B$  tiene un elemento típico  $a_{ij} + b_{ij}$ , mientras que  $A - B$  tiene el elemento típico  $a_{ij} - b_{ij}$ . Sigue fácilmente que  $A + B = B + A$  y que  $(A + B)' = A' + B'$ .

Una matriz  $A$  de dimensiones  $n \times k$  y una matriz  $B$  de dimensiones  $k \times m$  pueden ser multiplicadas para producir una matriz de dimensiones  $n \times m$ . Considerando el caso especial donde  $k = 1$ , entonces  $A = a'$  es un vector fila y  $B = b'$  es un vector columna. Entonces, se define:

$$AB = a'b = (a_1, a_2, \dots, a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = a_1b_1 + a_2b_2 + \dots + a_nb_n$$

Se denomina  $a'b$  como **producto interno** o producto punto de los vectores  $a$  y  $b$ . Note que  $a'b = b'a$ . Dos vectores son llamados ortogonales si  $a'b = 0$ . Para cualquier vector  $a$ , excepto el vector nulo, se tiene que  $a'a > 0$ . El producto exterior de un vector  $a$  es  $aa'$ , que es de dimensión  $n \times n$ .

Un caso especial surge para  $m = 1$ . En este caso  $A$  es una matriz  $n \times k$  y  $B = b$  es un vector de dimensión  $k$ . Entonces  $c = Ab$  es también un vector, pero de dimensión  $n$ . Sus elementos son:

$$c_i = a_{i1}b_1 + a_{i2}b_2 + \dots + a_{ik}b_k$$

que es el producto interno entre el vector obtenido de la fila  $i$  de  $A$  y el vector  $b$ .

Cuando  $m > 1$ ,  $B$  es una matriz y  $C = AB$  es una matriz de dimensión  $n \times m$  con elementos típicos

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{ik}b_{kj}$$

siendo el producto interno entre los vectores obtenidos de la fila  $i$  de  $A$  y la columna  $j$  de la matriz  $B$ . Note que este solo hace sentido si el número de columnas de  $A$  es igual al número de filas en  $B$ .

Considere el siguiente ejemplo:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 0 \end{pmatrix}$$

y

$$B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 5 \end{pmatrix}$$

y

$$AB = ?$$

Es importante notar que  $AB \neq BA$ . Incluso si  $AB$  existe,  $BA$  no puede estar definida, debido a que las dimensiones de  $B$  y  $A$  no son conformables. Si  $A$  es de dimensión  $n \times k$  y  $B$  es de dimensión  $k \times n$ , entonces  $AB$  existe y tienen dimensión  $n \times n$ , mientras que  $BA$  existe con dimensión  $k \times k$ . En el ejemplo anterior, se tiene que

$$BA = \begin{pmatrix} 9 & 12 & 3 \\ 19 & 26 & 9 \\ 20 & 25 & 0 \end{pmatrix}$$

Para la transpuesta de un producto de dos matrices, se tiene que

$$(AB)' = B'A'$$

De aquí (y  $(A')' = A$ ) sigue que ambas  $A'A$  y  $AA'$  existen y son simétricas. Finalmente, multiplicando un escalar y una matriz es lo mismo que multiplicar cada elemento en la matriz por este escalar. Esto es, para un escalar  $c$ ,  $cA$  tiene como elemento típico  $ca_{ij}$

## Propiedades de las matrices y vectores

Si se considera un número de vectores  $a_1$  a  $a_k$ , se puede tener una **combinación lineal** de estos vectores. Con pesos los escalares  $c_1, \dots, c_k$  este produce el vector  $c_1a_1 + c_2a_2 + \dots + c_ka_k$ , que puede ser escrito como  $Ac$ , donde, como antes  $A = [a_1, \dots, a_k]$  y  $c = (c_1, \dots, c_k)'$ .

Un conjunto de vectores es **linealmente dependiente** si cualquier de los vectores pueden ser escritos como una combinación lineal de los otros. Esto es, si existen valores para  $c_1, \dots, c_k$ , no todos cero, tal que  $c_1a_1 + c_2a_2 + \dots + c_ka_k = 0$  (el vector nulo). Igualmente, un conjunto de vectores es linealmente independiente si la solución a

$$c_1a_1 + c_2a_2 + \dots + c_ka_k = 0$$

es

$$c_1 = c_2 = \dots = c_k = 0$$

Esto es, si la única solución de  $Ac = 0$  es  $c = 0$ .

Si se consideran todos los vectores posibles que pueden tenerse como combinación lineal de los vectores  $a_1, \dots, a_k$ , estos vectores forman un **espacio vectorial**.

Si los vectores  $a_1, \dots, a_k$  son linealmente dependientes, se puede reducir el número de vectores sin cambiar el espacio vectorial. El número mínimo de vectores necesarios para generar un espacio vectorial es llamado **dimensión** de ese espacio. En este sentido se puede definir el **espacio columna** de una matriz como el espacio generado por sus columnas, y el **rango columna** de una matriz como la dimensión de su espacio columna. Es claro, que el rango columna nunca puede exceder el número de columnas. Una matriz es de rango columna completo si el rango columna es igual al número de columnas. El **rango fila** de una matriz es la dimensión del espacio generado por las filas de una matriz. En general, se sostiene que el rango fila y el rango columna de una matriz son iguales, de tal manera que inambiguamente define el rango de una matriz. Note que esto no implica que una matriz que es de rango columna completo es automáticamente de rango fila completo (esto solo se cumple si la matriz es cuadrada).

Un resultado útil en el análisis de regresión para cualquier  $A$  es:

$$\text{rango}(A) = \text{rango}(A'A) = \text{rango}(AA')$$

## Matrices Inversas

Una matriz  $B$ , si existe, es la inversa de la matriz  $A$  si  $AB = I$  y  $BA = I$ . Un requerimiento necesario para esto es que  $A$  sea una matriz cuadrada y tenga rango completo. En este caso es llamada **invertible** o **no singular**. En ese sentido, se puede definir  $B = A^{-1}$ ,

$$AA^{-1} = I$$

y

$$A^{-1}A = I$$

Note que la definición implica que  $A = B^{-1}$ . Entonces, se tiene que  $(A^{-1})^{-1} = A$ . Si  $A^{-1}$  no existe, decimos que  $A$  es **singular**. Analíticamente, la inversa de una matriz diagonal y la inversa de una matriz  $2 \times 2$  se obtienen fácilmente. Por ejemplo,

$$\begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}^{-1} = \begin{pmatrix} a_{11}^{-1} & 0 & 0 \\ 0 & a_{22}^{-1} & 0 \\ 0 & 0 & a_{33}^{-1} \end{pmatrix}$$

y

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Si  $a_{11}a_{22} - a_{12}a_{21} = 0$ , la matriz  $A$   $2 \times 2$  es singular: sus columnas son linealmente dependientes, y por lo tanto sus filas. Se denomina a  $a_{11}a_{22} - a_{12}a_{21}$  el **determinante** de esta matriz  $2 \times 2$ .

Suponga que se requiere resolver  $Ac = d$  para un  $A$  y  $d$  datos, donde  $A$  es de dimensión  $n \times n$  y ambos  $c$  y  $d$  son vectores  $n$ -dimensionales. Este es un sistema de  $n$  ecuaciones lineales con  $n$  incógnitas. Si  $A^{-1}$  existe, se puede escribir:

$$A^{-1}Ac = c = A^{-1}d$$

para obtener la solución. Si  $A$  es no invertible, el sistema de ecuaciones lineales tiene dependencias lineales. Hay dos posibilidades: más de un vector satisface  $Ac = d$ , de tal manera que no existe solución única o las ecuaciones son inconsistentes, por tanto no existe solución al sistema. Si  $d$  es el vector nulo, solo la primera posibilidad es posible.

Fácilmente se puede derivar que:

$$(A^{-1})' = (A')^{-1}$$

y que

$$(AB)^{-1} = B^{-1}A^{-1}$$

(asumiendo que ambas existan).

## Matrices idempotentes.

Un caso especial es el de matrices simétricas e idempotentes. Una matriz  $P$  es **simétrica** si  $P' = P$  e **idempotente** si  $PP = P$ . Una matriz simétrica idempotente  $P$  tiene la interpretación de una **matriz de proyección**. Esto significa que el vector de proyección  $Px$  está en el espacio columna de  $P$ , mientras que el vector residual  $x - Px$  es ortogonal a cualquier vector en el espacio columna de  $P$ .

Una matriz de proyección que proyecte el espacio columna de una matriz  $A$  puede ser construido como  $P = A(A'A)^{-1}A'$ . Claramente, esta matriz es simétrica e idempotente. Proyectando dos veces sobre el mismo espacio, no altera el resultado, de tal manera que se tiene que  $PPx = P$ , que resulta directamente. El residuo de la proyección es  $x - Px = (I - A(A'A)^{-1}A')x$  tal que  $M = I - A(A'A)^{-1}A'$  es también una matriz de proyección con  $MP = PM = 0$  y  $MM = M = M'$ . Entonces, los vectores  $Mx$  y  $Px$  son ortogonales.

Una matriz de proyecciones interesante, que se usará más adelante, es  $Q = I - (1/n)u'$ . Donde  $u'$  es una matriz de unos. Los elementos de la diagonal de esta matriz son  $1 - 1/n$ , y los elementos fuera de la diagonal son  $-1/n$ . Ahora,  $Qx$  es un vector que contiene las desviaciones de  $x$  de su media. Un vector de medias es producido por la matriz de transformación  $P = (1/n)u'$ . Note que  $PP = P$  y  $QP = 0$ .

La única matriz de proyección que no es singular es la matriz identidad. Las demás matrices de proyección son singulares, cada una con el rango igual a la dimensión del espacio sobre la que se hará la proyección.

## Valores y vectores propios

Sea  $A$  una matriz simétrica  $n \times n$ . Considere el siguiente problema de encontrar combinaciones de un vector  $c$  (otro diferente al vector nulo) y un escalar  $\lambda$  que satisfacen:

$$Ac = \lambda c$$

En general, hay  $n$  soluciones  $\lambda_1, \lambda_2, \dots, \lambda_n$  llamados **valores propios** (raíces características) de  $A$ , correspondientes a  $n$  vectores  $c_1, c_2, \dots, c_n$  llamados **vectores propios** (vectores característicos). Si  $c_1$  es una solución, entonces también es cierto que  $kc_1$  es una solución para cualquier constante  $k$ , de tal manera que los vectores propios están definidos hasta una constante. Los vectores propios de una matriz simétrica son ortogonales, esto es,  $c'_i c_j = 0$  para todo  $i \neq j$ .

Si un valor propio es cero, el vector correspondiente  $c$  satisface  $Ac = 0$  que implica que  $A$  no es de rango completo y por lo tanto singular. Entonces, una matriz singular tiene al menos un valor propio igual a cero. En general, el rango de una matriz simétrica corresponde al número de valores propios diferentes de cero.

Una matriz simétrica es llamada **positiva definida** si todos sus valores propios son positivos. Es llamada **positiva semidefinida** si todos sus valores propios son no negativos. Una matriz positiva definida es invertible. Si  $A$  es positiva definida, se tiene que para cualquier vector de  $x$  (diferente al vector nulo) que

$$x'Ax > 0$$

La razón es que cualquier vector  $x$  puede ser escrito como una combinación lineal de los vectores propios como  $x = d_1 c_1 + \dots + d_n c_n$  para los escalares  $d_1, \dots, d_n$ , y se puede escribir

$$x'Ax = (d_1 c_1 + \dots + d_n c_n)' A (d_1 c_1 + \dots + d_n c_n) = \lambda_1 d_1^2 c_1' c_1 + \dots + \lambda_n d_n^2 c_n' c_n > 0$$

De manera similar, para una matriz  $A$  positiva semi-definida, se tiene para cualquier vector  $x$

$$x'Ax \geq 0$$

El **determinante** de una matriz simétrica es igual al producto de sus  $n$  valores propios. El determinante de una matriz positiva definida es positivo. Una matriz simétrica es singular si el determinante es cero (eso es, si uno de los valores propios es cero).

## Diferenciación

Sea  $x$  un vector columna  $n$ -dimensional. Si  $c$  es también un vector columna  $n$ -dimensional,  $c'x$  es un escalar. Considere  $c'x$  como una función del vector  $x$ . Entonces, se puede considerar el vector de derivada de  $c'x$  con respecto a cada uno de los elementos en  $x$ , esto es,

$$\frac{\partial c'x}{\partial x} = c$$

Este es un vector columna con  $n$  derivadas, donde el elemento típico es  $c_i$ . De manera más general, para una función vectorial  $Ax$  (donde  $A$  es una matriz) se tiene:

$$\frac{\partial Ax}{\partial x} = A'$$

El elemento en la columna  $i$ , fila  $j$  de la matriz es la derivada del elemento  $j$  en la función  $Ax$  respecto a  $x_i$ .

Más aun,

$$\frac{\partial x'Ax}{\partial x} = 2Ax$$

para una matriz  $A$  simétrica. Si  $A$  no es simétrica, se tiene que

$$\frac{\partial x'Ax}{\partial x} = (A + A')x$$

Todos estos resultados resultan de coleccionar los resultados de una diferenciación elemento por elemento.

## Algunas manipulaciones es de mínimos cuadrados:

Sea  $x_i = (x_{i1}, x_{i2}, \dots, x_{iK})'$  con  $x_{i1} \equiv 1$  y  $\beta = (\beta_1, \beta_2, \dots, \beta_K)'$ . Entonces,

$$x_i' \beta = \beta_1 + \beta_2 x_{i2} + \dots + \beta_K x_{iK}$$

La matriz

$$\begin{aligned} \sum_{i=1}^N x_i x_i' &= \sum_{i=1}^N \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iK} \end{pmatrix} (x_{i1}, x_{i2}, \dots, x_{iK}) \\ &= \begin{pmatrix} \sum_{i=1}^N x_{i1}^2 & \sum_{i=1}^N x_{i2} x_{i1} & \cdots & \sum_{i=1}^N x_{iK} x_{i1} \\ \vdots & \sum_{i=1}^N x_{i2}^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ \sum_{i=1}^N x_{i1} x_{iK} & & \cdots & \sum_{i=1}^N x_{iK}^2 \end{pmatrix} \end{aligned}$$

es una matriz  $K \times K$  simétrica que contiene la suma de cuadrados y los productos cruzados. El vector:

$$\sum_{i=1}^N x_i y_i = \begin{pmatrix} \sum_{i=1}^N x_{i1} y_i \\ \sum_{i=1}^N x_{i2} y_i \\ \vdots \\ \sum_{i=1}^N x_{iK} y_i \end{pmatrix}$$

tiene tamaño  $K$ , de tal manera que el sistema

$$\left( \sum_{i=1}^N x_i x_i' \right) b = \sum_{i=1}^N x_i y_i$$

es un sistema de  $K$  ecuaciones con  $K$  incógnitas (en  $b$ ). Si  $\sum_{i=1}^N x_i x_i'$  es invertible, una solución única existe. La invertibilidad requiere que  $\sum_{i=1}^N x_i x_i'$  sea de rango completo. Si no es de rango completo, existe un vector diferente de cero y de dimensión  $K$ , tal que  $x_i' c = 0$  para cada  $i$ , y existe una relación lineal entre las columnas/filas de la matriz  $\sum_{i=1}^N x_i x_i'$ .

En notación matricial, la matriz  $N \times K$   $X$  es definida como:

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1K} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NK} \end{pmatrix}$$

y  $y = (y_1, y_2, \dots, y_N)'$ . De ahí se puede verificar que

$$X'X = \sum_{i=1}^N x_i x_i'$$

y

$$X'y = \sum_{i=1}^N x_i y_i$$